⑩日本国特許庁(JP)

①特許出願公開

母公開特許公報(A) 平3-113563

@Int, Cl.* G 06 F 15/16 微別記号 380 Z

庁内整理番号 6945-5B @公陽 平成3年(1991)5月14日

· 10/10 3/6 U Z 6945-91

審査請求 未請求 請求項の数 16 (全18頁)

の発明の名称 マルチプロセッサスケジューリング方法

②特 願 平1-250532

砂岩 明 者 山口 停一期 茨城県日立市久窓町4026番地 株式会社日立製作所日立研 安府内

勿免 明 者 齊 藥 雅 彦 茨城県日立市久慈町4028番地 株式会社日立製作所日立研 作所内

個第 明 者 小 林 芳 街 茨城県日立市久慈町4026番地 株式会社日立製作所日立研 完所内

の出 顧 人 株式会社日立製作所 の代 理 人 弁理士 秋本 正実 最終頁に続く 東京都千代田区神田駿河台4丁目6番地

明報

/ . 意明の名称

マルチプロセッサスケジューリング方法

- は、特許請求の難問
 - 1.複数のプロセッサを共偏し、鉄模数のプロセ ッサが、単一のオペレーティングシステムによ って創御されるマルチプロセッサシステムにお いて、プロセスを管理するプロセス管理テープ ルに使い、上記オペシーティングシステムのブ ロセススケジューラが、プロセスに何れかのブ ロセッツを割り掛てるプロセススケジューリン グ方式であって、上記プロセス管理テーブルド 着棋を許すグループ離別子を記述し、味グルー プ機器子が同一な該プロセスの情報を耐像する グループ管理テーブルを設け、該グループ管理 テーブルにダループ内のプロセスを契行するプ ロセッサを提示するプロセッサ割り当て情報を 記述し、上記プロセススケジューラは、次に夾 行すべく差択したプロセスについて、上記プロ セス管理テーブルを参照して、当試プロセスの

- f -

グループを刺激し、値グループのグループ管理 テーブルのプロセッサ制り当て情報により指定 されたプロセッサ上ではプロセスを動作させる ことにより、同一のグループに属するプロセス は、同一のプロセッサで実行することを特徴と するマルチプロセッサスケジューリング方法。

- 3. 上記グループ智楽テーブルに彼グループの何 れのプロセスも実行されていない特徴、限ちグ

铃朗平 3-113563(2)

ループ実行的も時間を記憶し、上記プロセスス ケジューラが、次に災行すべく選択したプロセ スについて、上記プロセス管恵テーブルを参照 して、最終プロセスのグループを判別し、肢グ ループのグループ智楽テーブルのグループ実行 終ち時間がある領域上の時には、該グループ智 **塩テーブルのプロセッサ割り当て貸租を変更で** さるようにしたことを約数とするマルチプロセ ッサスケジューリング方法。

4.上記プロセス哲麿チーブルに減プロセスの実 行称ち時間とはプロセスが前回実行されたプロ セッサの裁割子を記憶し、上記プロセススケジ ューラが、次に実行すべく選択したプロセスに ついて、上記プロセス管理テーブルを参照して、 鎮持ら時間がある性以下の時には、鉄プロセス を前四尖打したプロセッタか、グループ管弧ナ ーブルのプロセッサ割り当て情報が混乐するプ ロセッサで実行し、ある龍以上の時には、上配 以外のプロセッサで粉作できるようにしたこと を特徴とする競戏項1又は2のマルチプロセッ

サスケジューリング方法。

- 6. 鉄プロセスの実行特ち時期の代わりに鉄プロ セスのプロセッヤへの割り占てが位船された国 数を用いたことを特徴とする結束項2又は8又 はものマルチプロセッサスケジューリング方法。
- 8. 鉄グループの何れのプロセスも実行されてい ない時間、即ちグループ実行券も時間の代わり に該プロセスのプロセッサへの割り当てが抵絶 された回路を貼いたことを特徴する請求項3又 は5のマルチプロセッサスケジューリング方法。
- 7.複数のプロセッサを具備し、関模数のプロセ ッサが、単一のオペレーティングシステムによ って制御されるマルチプロセッサシステムにお いて、実行しているプロセスを切り替える必要 が生じたプロセッサが、プロセスを替退するプ ロセス管理テーブルに従い、上記オペレーティ ングシステムのプロセススケジューラを実行し、 鉄プロセッサが次に尖行するプロセスを決定す るプロセススケージューリング方式であって、 上記プロセス管道テーブルに負担を許すグルー

. 4 -

プ策別子を記述し、該グループ機制子が同一な 鉄プロセスの錯裂を記録するグループ管理テー ブルも設け、鉄グループ管理テーブルにグルー プ内のプロセスを実行するプロセッサを知示す るプロセッサ割り当て情報を記述し、上記プロ セススケジューラは、実行可能プロセスの中か 6. 上記プロセス管道テーブルのグループ数別 子が指示する、はグループ管理テーブルのプロ セッサ前り組てĠ根が設プロセススケジューラ を共行しているプロセッサを俗示しているプロ セスを次に実行することを特徴とするマルチブ ロセッサスケジューリング方也。

- 8. 上記プロセス製廻テーブルに味プロセスの笑 行待ち時間を記憶し、上部プロセススケジュー ラが、次に表行するプロセスを選択する際に上 船プロセス管理テーブルを参照して、実行符ち 時間がある値以上の時には、ほプロセスを実行 可能とすることを影散とする間求収7のマルチ プロセッサスケジューリング方払。
- 9.上記グループ管理テーブルに欧グループの何

れのプロセスも実行されていない時間、即ちグ ループ実行持ち時間を記律し、上記プロセスス ケジューラが、次に実行するプロセスを表訳す る色に、上記プロセス智楽テーブルを参展して、 当該プロセスのグループを刺削し、弦グループ のグループ管道テーブルのグループ表行のちゅ 関がある領以上の時には、数グループ管理テー ブルのプロセッサ割り当て情報を寂寞して、核 プロセスを実行可能とすることを特徴とする話 **永項7又は8のマルチプロセッサスケジューリ**

10、上記プロセス管理テーブルにはプロセスの尖 行為ち時間と数プロセスが前回条行されたプロ セッサの識別子を記憶し、上記プロセススケジ ューラが、次に数行するプロセスを選択する際 に、上記プロセス管理テーブルを参照して、食 待ち時間がある質以下の時には、咳プロセスを 前回実行したプロセッサか、グループ管理テー プルのプロセッサ割り当て情報が相景するプロ セッサが眩プロセススケジューラを実行してい

-426--

特間平 3-113569(3)

るプロセッサと同一である場合にのみ実行可能 とし、載待ち時間がある他以上の時には、実行 可能とするこ、 お迷とする請求項7のマルチ プロセッテスケジューリング対象。

- 11. 放プロセスの変行的ち時間の代わりに放プロセスのプロセッサへの割り出てが揺続された回数を用いたことを特徴とする建来項8又は9又は10のマルチプロセッサスかジューリング方法。
- 12. はグループの何れのプロセスも実行されていない時間、知ちグルブ実行符も時間の代わりにはプロセスのプロセッサへの割り出てが包轄された即数を用いたことを特徴とする情報項8又は6のマルチプロセッサスケジューリング方法。
- 13. 複数のプロセッサを其個し、旅稼数のプロセッサが、ボーのオペレーティングシステムによって制御されるマルチプロセッサシステムにおいて、プロセスを管理するプロセス管理テーブルに従い、上記オペレーティングシステムのプロセッサを割り高でるプロセススケジューリン

- 14、プロセス割り当て情報のコードの一つとして、 放プロセスが何れのプロセッサでも実行可能で あることを示すコードを設けたことを特徴とす る辞求者1~13のいずれか1つに包穀のマルチ プロセッサスケジューリング方欲。
- 16. プロセス管理テーブル、グループ管理テーブルの何れか、または、両方にはプロセスが特定のプロセッサでのみ実行可能であることを数す

- 7 -

フラグを設けたことを特徴とする静水項1~14 のいずれか1つに記載のマルチプロセッサスケ ジューリング方法。

- 16、安原を共有するプロセスに対して、同一のグループ機関子を持たせることにより、証責係の 抑他制御を簡単化したことを特徴とする簡単項 1~15のいずれか1つに記載のマルチプロセッ サスケジューリング方法。
- 1. 発明の詳細な説明

【産業上の利用分野】

複数のプロセッサから成るマルチプロセッサシステムで複数のプロセスを並行に動作させるマルチプロセッサ用オペレーディングシステムにおいて、特に、高度システムを修築する場合、あるいはシングルプロセッサルに作成されたプログラムをマルチプロセッサシステムで、なんら製更なく動作させる場合に分語なプロセススケジューリンクオ#に関する。

【従来の技術】

ែ製のマイクロプロセッサを使用 したマルチマ

. g .

イクロプロセッサの開発が進んである。特別語55 - | 27261分、マルチプロセッサシステムはこのー 例であり、推動のプロセッサがバスにより始合さ れたシステム構成を示している。マルチプロセッ サ用オペレーティングシステムについても 盛んに 如尭が行なわれている。マルチプロセッサ用オペ シーティングシステムにおいてはシングルプロセ ッサシステムとは異なり、複数のプロセスが異な るプロセッサ上で回時に実行できるようにする必 寒があり、このために、プロセスをどのプロセッ ひに割り当てるかを決定する方策を持つ必要があ る。この場合、最も一般的に行なわれているのは プロセッサの利用効率を長大限にするために、食 荷の最も少ないプロセッサに割り当てるという方 様である。この方象は一般的に実現が容易である ため、広く採用されている。というのも、雌々の プロセッサ上で動作するオペレーティングシステ ムのプロセススケジューラボぞれぞれシングルブ ロセッサシステムの毎合と同様に次に実行すべき プロセスをプロセスの研究制度に従って頭視する。

--427--

特別平 3-113563(4)

また、プロセスをグループとして管理して、スケジューリングする方法について、「エクスペアリンス・ユーズィング・マルチープロセッサ・レステムズーAスティダス・レポート」(*Biper-Lonce Using Multi-processor Systems-A Status Report*)(ACM Computing Surveys, 12巻2号p121~165, 1989)にグルー

プスケジューリング (group acheduling) として 配されている。これは、位列処理可能なプロセス をグループとして、スケジューリングするとをに はこれらのプロセスをできるかぎり同時に異なる プロセッサで実行するようにする。

[発明が無終しようとする鉄紙]

複数のプロセスをグループとし、同一グループ に属するプロセスをできるがきり同一のプロセッ やで実行する機能を有するプロセススケジューラ が好ましい。

ここでプロセスをグループ化するのに以下のようなユーザ要求およびハードウェアならびにオペレーティングシステム鉄路上の要求があるためである(荷、このグループとは同じデータを扱うプロセスの集まりである)。

共有バス、失有メモリを使ったマルチプロセッサでは、バスの負荷を軽くするため、各プロセッサがキャッシュメモリを押つ。この様なシステムでは、キャッシュ関でデータの一貫性を単純するために一つのプロセッサがキャッシュ上のデータ

- 11 -

を更新すると、それと同じデータを持っているキャッシュに更報した結構を収透する。お互いに多くのデータを共有しているプロセスを異なるプロセッサで実行するとキャッシュ間のデータを送の回数が増大し、キャッシュの性能が発揮されず、バスの気勢が高くなってしまう。使来のマルチプロセッサシステムでは、必ずしもこの問題点に関して十分な配慮がなされていなかった。

また、シングルプロセッサ用に開発された戦存のソフトウェアは、マルチプロセッサシステム上においても何等変更を加えずに共行支では、増知のように、複数のプロセスの整合が発生しつる。この問題を解決するため、通常、共有等では、増加のように、複数のプロセスの整合、共有等では、プロスする路にセマフォアメカニズム等によりよりも処理時間の類にしたり、セマフォアメカニズムよりも処理時間の類にいずな字の手段をよりを上げて前り込み鍼止にする字の手段を

- 12 -

はじる場合が多い。これらの手段は、プロセッサ 1 包上でもクロに見ればプロセッサ上の処理は必 ず遅次的に行なわれるという事実を利用している。 現実には、共有データに対する卵硫制料を性態上 の現由により行なっていないソフトウェアも多い。 このため既存のソフトウェアをその役マルチプロ セッサレステムで実行させた場合、必ずしも正常 な動作は領質されないという問題があった。

並来のgroup schedulingは、グループとなっているプロセッサを異なるプロセッサで動作をせているため、放剤処理には著しているが、上記のような問題には対応できなかった。

尚、本願に関係する免職は、同一出版人になる特願は1-60091号「マルチプロセッサ用オペレーティングシステム」がある。この免職はプロセス管理テーブルを持ち、このテーブルにプロセス方式プロセッサ名を記録しておき、他プロセッサ支配下のプロセスを一定条件のもとに検取りして自己の管理において処理を行なわせることとしている。この免額には、前回使用プロセッサの記録

特期平 3-113563(5)

やグループほのプロセス管理の記載はない。

本行明の目的は、マルチプロセッサシステムにおいて、プロセスをグループ化して管理し、同一グループ内のプロセスを何ープロセッサで実行することにより、キャッシュメモリの性能を向上させ、バスの負債を保険できるプロセススケジューリングが独を提供することにある。

本発明の間の目的は、既存の単一プロセッサシステム用のソフトウェアを、なんら変更させずその食ま利用できるプロセススケジューリング方法を提供することにある。

本売明の更に別の目的は、グループとなったプロセスを同一プロセッサで実行する際にも、プロセッサ間で仕事の負荷を均等に扱つことがでまるプロセススケジューリング方法を提供することにある。

『雑組を解決するための手段】

上記目的を連抜するために、本発明によるプロ セススケジューリング方法は、複数のプロセッサ を具督し、上記複数のプロセッサが、単一のオペ

. 15 -

特ち時間を記憶して、上記プロセススケジューラが、次に実行すべく選択したプロセスについて、上記プロセスは世理テーブルを参照して、実行特も時間がある値以上の時には、任意のプロマッドには、数プロセス管理テーブルを参照して、当該プロセスでは、放了ロセッサを判別し、放びループのグループを認って、対策により行れています。 されたプロセッサ上で該プロセスを動作させるようにしてもよい。

さらに上記グループ管理テーブルに鉄グループの何れのプロセスも実行されていない時間、即ちグループ実行的ち時間を配復し、上記ブロセススケジューラが、次に実行すべく選択したプロセスについて、上記プロセス管理テーブルを参照して、当試プロセスのグループを報期し、鉄グループのグループ管理テーブルのグループ管理テーブルのブループ管理テーブルのブロセッサ前り当て情報を変更できるようにしてもよい。

レーティングシステムによって刺棘されるマルテ プロセッサシステムにおいて、プロセスを管理す るプロセス管理テーブルに使い、上記オペレーテ ィングシステムのプロセススケジューラが、プロ セスに何れかのプロセッサを割り当てるプロセス スケジューリング方式であって、上記プロセス管 選チーブルに営機を許すグループ限別子を耐速し、 彼グループ側別子が同一な様プロセスの情報を記 後するグループ管理テーブルを載せ、彼グループ 管理テーブルにグループ内のプロセスを実行する プロセッサを組みするプロセッサ前り背で特保を 記述し、上記プロセススケジューラは、次に英行 すべく選択したプロセスについて、上記プロセス **智選テーブルを参照して、岩跡プロセスのグルー** プも判別し、妹グループのグループ管理テーブル のプロセッサ割り留て倒視により指定されたプロ セッサ上ではプロセスを動作させることにより、 同一のグループに属するプロセスは、同一のプロ セッサで実行するようにしたものである。

上記プロセス管理テーブルにはプロセスの実行

· 16 -

さらに、上記プロセス管理テーブルに該プロセスで取り付も時間と疑プロセスが簡単行されたプロセッサの預別子を配信し、上記プロセススケソューラが、次に実行すべく表択したを参照して、上記プロセスを開いて、上記プロセッサがには、放プロセスを節回式行したプロセッサか、グループ管理テーブルのプロセッサが到りまするが収入するプロセッサで実行し、ある他以上の時には、上記以外のプロセッサで動作できるようにしてもよい。

さらに、上記プロセスの実行特を時間の代わり に該プロセスのプロセッサへの割り当てが拒絶さ れた回数を用いてもよい。

さらに、上記グループの何れのプロセスも実行されていない時間、即ちグループ実行待ち時間の 代わりに繋ブロセスのプロセッサへの割り当てが 拒絶された回数を頂いた。

本発明の他のプロセススケジューリング方法と して、複数のプロセッサを共働し、該複数のプロ セッサが、単一のオペレーティングシステムによ

铃脚平 3-113563(6)

って創倒されるマルチプロセッサンステムにおい て、実行しているプロセスを切り替える必覧が生 じたプロセッサが、プロセスを管理するプロセス 管道テーブルに従い、上記オペレーティングシス テムのプロセススケジューラを実行し、放プロセ ソサが次に尖行するプロセスを決定するプロセス スケージューリング方式であって、上記プロセス 世雄ナーブルに重複を許すグループ撤別子を配途 し、紋グループ像別子が同一な孩プロセスの情報 も記載するグループ管理テーブルを取け、鉄グル ープ管理テープルにグループ内のプロセスを実行 するプロセッサを超示するプロセッサ割り当て様 役を配迹し、上記プロセススケジューラは、実材 可能プロセスの中から、上記プロセス管理テープ ルのグループ機別子が投示する。紋グループ管理 テーブルのプロセッサ部り誰で情報が抜プロセス スケジェーラを実行しているプロセッサを抱示し ているプロセスを次に気行するようにしてもよい。 上記プロセス管理テーブルに改プロセスの実行

次に設行するプロセスを選択する際に上記プロセスを選択するがはいた。 大管環ケーブルを参照して、実行符も時間がある は以上の時には、はプロセスを実行可能ととして

をもに、上記グループを選字ーブルにはグループの何れのプロセスも実行されていない時間、即ちグループ実行符も時間を記録し、上記プロセスを選択する際に、上記プロセスを選手ーブルを参照して、当該プロセスのグループを判別し、該グループのガループで選子ーブルのグループで選子ーブルのプレープを選び上の時には、該グループを選びーブルのプレープを選び上の時には、該グループを選びーブルのプレープを選び上の時には、該グループを選び上のよい。

さらに、上記プロセス管理テーブルに験プロセスの実行符も時間とはプロセスが前四実行されたプロセッサの裁別子を記憶し、上記プロセススケジューラが、次に実行するプロセスを選択する際に、上記プロセス管理テーブルを参照して、條件も時間がある低級下の時には、微プロセスを前回

- 19 -

待ち時悔を記憶し、上記プロセススケジューラが。

. 28 .

実行したプロセッサか、グループ管電テーブルの プロセッサ割り当て管根が指示するプロセッサが 該プロセススケジューラを実行しているプロセッ サと何一である場合にのみ実行可能とし、鉄約ち 時間がある彼以上の時には、実行可能としてもよ い。

さらに、上記プロセスの内行物も時間の代わり に該プロセスのプロセッサへの割り過でが範絶だ れた侵敗を用いてもよい。

さらに、上記グループの何れのプロセスも実行されていない時間、即もグルプ実行終も時間の代わりにはプロセスのプロセッサへの割り当てが想象された回数を用いた。

本発明の他のプロセススケジューリング方式として、複数のプロセッサを具備し、放複数のプロセッサを具備し、放複数のプロセッサが、単一のオペレーティングシステムにおいて、プロセスを替還するプロセス管理テーブルに従い、上記サペレーティングシステムのプロセススケジューラが、プロセスに何れかのプロセッサ

を割り当てるプロセススケジューリング方式であって、上記プロセス智雄テーブルに上記複数のプロセス智雄テーブルに上記複数のプロセッサを記録と、我グループ機関子が同一なかけ、など、アーブルにグループとない。大きなでは、大きなでは、大きなでは、大きなでは、上記プロセッサを担保するプロセッサを担保を設け、上記プロセスを関ラーブルのプロセッサのでは、上記プロセッサを担保テーブルのグループを関サーズが対し、上記プローブを直で、アーブルのグループを関サーズを関サーズを関サーズを関サーズを関サーズを関サーズを表示を表示という。

作用】

本書明のプロセススケジューリング方法は、複数のプロセッサからなるマルチプロセッサンステムにおいて、複数のプロセッサのいずれでも実行可能なプロセスについて、あえてそのプロセスをグループ化しそのグループを実行させるプロセッサを特定のプロセッサに限定しようとするものである。

特別平 3-113563(7)

共存データを多く持つプロセスが異なるプロセ ッサで表行されると、キャッシュ一致化のための 遺伝が影響に発生しキャッシュ性性を低下をせる。 女奇明では、そのようなプロセスをグループ化し、 即一のプロセッサで実行できるので、キャッシュ 一放化のための通信が不受となり、性能を向上さ せることができる。ちらに、グループを特定のプ ロセッサに抜行させることによりプロセッサ間の 負債が片寄ることがないように、一定時間以上次 行されずに放置されているプロセスについては、 いずれのプロセッサでも尖行できるようにしてお り、食药分散が可能となっている。ただし、この 仮能はプロセス固定フラグを立てることにより味 とすることもできる。

また、本務項によって、提他副仰を必要とする プロセス群を何一のプロセッサで動作させること が可能になり、共省メモリに対する特色制御を跨 慮していない姓来の承一プロセッツ用の抵存のプ ロセスを禁止することなく使用しても正常な動作 を保証することができる。この事は、他のプロセ

ッサの不使用を意味するものではなく、お兄プロ セス条以外のプロセスについては(このような症 的を改けずに、任意のプロセッサで実行可憶とし てマルチプロセッサの独的を有効に利用すること ができる。

【寒飲粉】

以下、本発明の一実施別を懲る懲を用いて説明 する。毎1団において、13-1~13-2はプロセ ッサ、12-1~12-8はプロセス、15-1~15-3はグループである。

プロセッサ13-1~13-8は、プロセス12-1 ~12-8を切り替えながら実行する。プロセス12 - 1 ~12 - 3 はグループ18 - 1 として、プロセス 12-4と12-5はグループ15-2として、プロセ ス12-6~12-8はグループ15-3として、まと められている。プロセスをグループにする方絵の 視点は、ユーザによってなされるが、コンパイラ やオペレーティングシステムが自動的に行なって **も良い。**

プロセスを実行するプロセッサはグループ伝に

- 23 -

決まっており、関一のグループに属するプロセス は、同一のプロセッサで実行するようにする。烙 1因ではグループの数とプロセッサの数が等しく なっているが、これらの数は、特に一枚している 必要はなく、どちらが多くても少なくても良い。 本島明によるシステム養成を好る図に示す。烙 2回において、13-1~13-3はプロセッサ、11 - 1~11・8はプロセッサ番号、12・1~12-8 はプロセス、)4-1はタイマー、4体技能パス、 5 は共有メモリ、31は06プログラム、311はス ケジューラ、32・1~32・4はプロセス管理テー ブル、321-1~321-4はプロセッサ意助子、 312-1~822-4は低先度、323-1~328-4は 実行特も時間、324 - 1 ~824 - 4 はカウンタ、 325・1~826・4はプロセス固定フラグ、826-1 ~326-4はグループ級別子、33-1~33-2は グループ管理テーブル、3\$L- 1~\$31-2はプロ セッサ級別子、832-1~332-2は契行的も時間、 889-1~838-8はカウンタ、384-1~384-2 はグループ囚定フラグである。

- 24 -

プロセッサ18 - 1 ~13 - 3 は、接続バス4によ り相互に、ならびに共有メモリらに独貌されてい る。各プロセッサはプロセッサ番号11-1~11-3をそれぞれ持っている。本実施例では、プロセ ッサ13・1~13・8はそれぞれプロセッサ番目 "1", "2", "3"を有している。このプロセッサ 着号の登定は、本茨族例ではハードウェアスイッ チにより行なっているが、他の方法として、ハー ドのROMにあらかじめ記憶しておく、立ち上げ 時にプログラムにより設定する等の力払も考えら

プロセッサ18・1~18・3のうち、一つのプロ セッサにはタイマー14~1があり、一定の時間お らにタイマー割り込みがそのプロセッサに入る。 各プロセスはプロセッサ上で実行される。共有 メモリSには、オペレーティングシステムのプロ グラム31、このプログラム31の一部であるプロセ ススケジューラ311、このオペレーティングシス テムが管理するテーブルの一部であるプロセス特 赵テーブル31(32・1~31・3)、グループ智道

特別平 3-113563(8)

テーブル33(83・1~33・1)が存在する。

プロセス管理テーブル3tには、以下のような役 役が置かれる。据1にそのプロセスを実行すべき プロセッサのプロセッサ券券を示すプロセッサ職 別子11 (321-1~321-4) である。実施例では プロセス管道テーブル32-1のプロセッサ裁別子 は"1"となっており、このプロセス管理テーブル に対応するプロセス0 (12・1) は、プロセッサ 着号ロー1に"1"を拾つ、プロセッサ13-1で爽 行すべきことを示している。プロセスを流行する プロセッサが未定のときは、このプロセッサ機別 子をりとしておく。また、虫行するプロセッサを 特に決めず、どのプロセッサでもそのプロセスを 実行できるようにしたい場合には、 プロセッサ酸 別子を会の値にしておく。

館2に、プロセスの実行展を決定する優先度 322 (322 - 1 ~322 - 4) である。この優先展の 数機が小さいほど、プロセスの優先度が高くなり、 他のプロセスに延先して実行されなければならな いことを示す。

. 27 .

るプロセッサ袋号が記されている。

第24、类疗特多時間322(322-1~322-2) である。この姿符符ち時間は、グループに属する プロセスのいずれもがプロセッサで実行されてお らず、いずれかが実行符も状態となっている時間 である。グループのプロセスが1つでも実行され るとこの実行符ち時間は0に戻される。

類3は、カウンタである。このカウンタはグル ープに属するプロセスの実行が組織された自教を 教えている。グループ内のプロセスが忠行された 時には0にクリアされる。

難4は、グループ国家フラグであり、このフラ グが立っているときには、そのグループを実行す。 るプロセッサは祖立となる。

節8団は、オペレーティングシステムのプログ ラム31の一節であるプロセススケジューラ311の アルゴリズムを示している。

まず、実行符ち状態となっているプロセスの中 から、低気異似に従い、次に実行するプロセスを 路沢する(511・1)、選択したプロセスのプロセ

. 29 -

男8は、実行役も時間328(323・1~328・4) である。この安行待ち時間は、そのプロセスが実 行券ち状態になっている時間を示す。プロセスが 、栄持されるシミの栄持得ち診職はうど寒される。

第4は、カウンタ324(324-1~324-4)で ある。このカウンタはプロセスの実行がお除され た四数を数えるカウンタである。

男5は、プロセス団定フラグ328 (328-1~326 - 4)である。ここには、プロセスをプロセッサ に固定にするかどうかの情報が入る。固定として おくとそのプロセスは、プロセッサ歳別于321に 怠されたプロセッサ以外では気行されない。

第8は、グループ設別子826(326・1~326-4) である。ここには、そのプロセスがどのグル ープに属しているかを示す情報が入る。グループ の無例子とグループ管理テーブルは1分1に対応

グループ管理テーブルには、以下の情報が配を れている。まず、毎1にプロセッサ限別子である。 これは、そのグループに属するプロセスを実行す

. 28 -

ス智酒テーブルより、グループ難別子325を読み だす。このグループ推別チよりグループ管道テー プルを検索する (811 - 2) 。

このグループ管理テーブル中のプロセッサ酸別 子53]を総みだし、この値により以下の処理を行 なう (311 - 8).

その低が正の低だったならば、スケジューラブ ログラム311を実行しているプロセッサは、自分 のプロセッサ番号を11(11・1~11・3) から続 みだし、その正のプロセッチ維別子331の値と比 校する(811-5)。一致していないときには、汝 に銀党献位が高いプロセスを選択し(3!!-1)。 一致しているときには選択したプロセスを自分の プロセッサに割り茹てる (311-6)。

グループ管理テーブルのプロセッサ散別子がO だったときには、そのプロセッサ識別子の値を自 分のプロセッサ番号 (11) の趙に変更して(3£1 -4)、遊択したプロセスを自分のプロセッサに割 45 C& (\$L1 - 6).

プロセッサ戦別子が失の世だったときには、無

特別平 3-113563(9)

条枠に透択したプロセスを合分のプロセッサに割 り当てる(\$11 - 6)。

第4 因は、オイ・ーティングシステムのプログ ラム31の一部であるプロセススケジューラ311の アルゴリズムの到の海的新を示している。

まず、実行等も状態となっているプロセスの中から、保免類位に従い、次に実行するプロセスを選択する (311-1)、選択したプロセスのプロセス管理テーブル31より、プロセッサ機関チ321を飲みだし、自分のプロセッサ機号11と比較する (311-7)。その値が等しかったときには、選択したプロセスを自プロセッサに割り当てる(311-6)。

等しくないときには、プロセス間窓フラグ325が立っているかどうか判定する (31(-8)。立っていた場合には、次に仮鬼眼位が高いプロセスを 題択し患す (311-1)。

立っていないときには、プロセス管理テーブル 32の実行待ち時間323がしまい彼を超えていない か悪べる (311 - 9)。

- 31 -

番号11に設定して(311-13, 311-14)、選択したプロセスを実行する(311-6)。

第5 関は、オペレーティングシステムのプログ うム31の一部であるプロセススケジューラ311の アルゴリズムの別の実践例を示している。

まず、実行特も状態となっているプロセスの中から、例先順位に従い、次に実行するプロセスを選択する(311-1)。通訊したプロセスのプロセス管理テーブル32より、プロセッサ機別子321を飲みだし、自分のプロセッサ機分引と比較する(311-7)。その値が等しかったときには、選択したプロセスを自プロセッサに割り出てる(311-6)。

答しくないときには、プロセス固定フラグ325が立っているかどうか利定する (311 - 8)。立っていた場合には、プロセス管理テーブル32のカウンタ324とグループ管理テーブル32のカウンタ333をインクリメントして (311・17, 311-18)、次に優先版位が高いプロセスを選択し直す(311-1)。

プロセス原党フラグが、立っていないときには、

超えている場合には、プロセス管式テーブル31 のプロセッサ課別子321を食プロセッサ音号に収 定して (311~14)、番択したプロセスを自分に割 り付ける (311~5)。

超えていない場合には、プロセスを選テーブル 32中のグループ酸別子326よりグループ管理テー ブルを検索し、その中のプロセッサ報別子が自分 のプロセッサ番号と等しいか調べる (311-10)。 等しい場合には、やはり、811-14を実行する。

等しくない場合には、グループが選テーブル33 中のグループ間定フラグ384を調べる (311-11)。 フラグが立っているときには、次に優先度の高い プロセスを削べる (311-1)。

フラグが立っていないときには、グループ管項テーブル33の実行時ち時間332がしせい値を超えていないが関べる。超えていない場合には、次に優先度の高いプロセスを耐べる(311-1)。

超えていた場合には、グループ管理テーブル83 のプロセッサ機関于331とプロセッサ管理テーブル82のプロセッサ機関于321を自分のプロセッサ

- 32 -

プロセス管理テーブル32のカウンタ325がしきい 値を組えていないか個人名(311 - 15)。

超えている場合には、プロセス管理テーブル32 のプロセッサ戦別平321を自プロセッサ番号に設 定して (311 - t4)、選択したプロセスを自分に割 り付ける (311 - 6)。

超えていない場合には、プロセス登越テーブル82中のグループ酸別子325よりブループ管理テーブルを検索し、その中のプロセッサ酸別子が自分のプロセッサ番号と等しいか飼べる(311-14)。 等しい場合には、やはり、311-14を実行する。

等しくない場合には、グループ管理テーブル33 中のグループ磁交フラグ334を構べる(311・11)。 フラグが立っているときには、311-17を実行す

フラグが立っていないときには、グループ管理 テーブル33のカウンタ333がしまい着を悠えてい ないか調べる。超えていない春台には、311・17 を笑行する。

超えていた場合には、グループ管理テーブル33。

特闘平 3-113563(10)

のプロセッサ機関子331とプロセッサ管理テーブル32のプロセッサ機関子321を自分のプロセッサ 番号11に反応して(211 - 12、311 - 14)、通気し をプロセスを実行する(311 - 6)。

係6回は、第3回、第4回、第5回の311-6 の選択したプロセスを含プロセッサに初り出てる 力物を辞載に記したものである。

まず、今まで央行していたプロセスのレジスタ を返避する(313-1)。これは、後にそのプロセ スに実行の服器が再び超ってきたときに、中耐し たところから再関できるようにするためである。

次に直視したプロセスのプロセス管理テーブル 31の実行終ち時間323を0にする(312・2)。

同様に選択したプロセスが属しているグループ のグループ智道テーブル32の次行符も時間332、 プロセス管理テーブル32のカウンタ324、グルー プ管理テーブルのカウンタ334も0にする(それ ぞれ、312-8,312-4,312-5)。

プロセスは、それぞれ独立なユーザ空間を持っ ているため、ユーザの空間を選択したプロセスの ものに張り替える (312-8)。

最後に、プロセスを節値中断したときに迅速しておいたレジスタの値を回旋する (312-7)。 通過するレジスタにはプログラムカウンタも含まれるため、前回中断した場所から処理が再図される。

第7回の315は、第1回14-1のタイマーがプロセッサに送る割り込みの処理ルーチンを示している。タイマー割り込みは10ma同隔で起こる。

タイマー割り込みが入るとプロセッサは、まず、 プロセス管理テーブル37を企影調べていき、実行 待ち状態となっているプロセスについては、その 実行給ち時数823をインクリメントする(319-1)。

次にグループ管理テーブル33を全部関べていき、 奥行中のプロセスを含まず。 奥行券ものプロセス を含むグループの終ち時間331をインクリメント する (313 - 2)。

第2回の本売頃のシステム構成図の別の実施例 も即6回に示す。第8回において、18-1~18-3はプロセッサ、11-1~11-3はプロセッサ等 号、12-1~12-6はプロセス、14-2はタイマ

- 35 -

ー、4は接続パス、6は共有メモリ、31は06プログラム、31:はスケジューラ、32・1~32・4はプロセス管理テーブル、321・1~321・4はプロセン管理テーブル、321・4はプロセン関盟テ、325・1~325・4はプロセス固定フラダ、326・1~326・4はグループ課別テ、327・1~327・4はCPU時間、328・1~328・4、329・1~320・4、330・1~330・4は研免度、33・1~33・2はグループ管理テーブル、331・1~331・2 なびアープログフラグである。

プロセッサ13-1~13-3は、接続パス4により相互に、ならびに共有メモリ5に接続されてい、る。各プロセッサはプロセッサ番号11-1~11-3をそれぞれ投っている。本央復例では、プロセッサ13-1~18-3はそれぞれプロセッサ番号*1*.*8"、*3"を有している。

プロセッサ13-1~13-8のうち、一つのプロセッサにはタイマー14-2があり、一定の時間おきにタイマー初り込みがそのプロセッサに入る。 実施例では10msおきと1mおきの2種類の割り込 . 36 -

みがみる.

各プロセスはプロセッサ上で実行される。共有 メモリ 5 には、オペレーティングシステムのプロ グラム31、このプログラム31の一部であるプロセ ススケジュー 9 314、このオペレーティングシス テムが甘地するテーブルの一部であるプロセスサ 班テーブル32(32 - 1 ~ 32 - 3)、グループ管理 テーブル33(33 - 1 ~ 33 - 2)が存在する。

プロセス管理テーブル32には、以下のような情報が置かれる。第1にそのプロセスを実行すべきプロセッサのプロセッサ番号を示すプロセッサ酸別子21(321-1~321-4)である。 突旋例ではプロセス管理テーブル32-1のプロセス特別が一大では、プロセスを受け、このプロセスを明テルに対応するプロセスの(12-1)にプロセッサの共和のときは、プロセッサが未定のときは、このプロセッサを予けている。 また、実行するプロセッサが未定のとき、実行するプロセッサが未定のときな、実行するプロセッサが未定のときな、実行するアロセッサをものプロセッサでもそのプロセスを

特開平 3~113563(11)

実行できるようにしたい場合には、プロセッサ種 割子を虫の低にしておく。

類2は、プロ1 (固定フラグ325 (323 - 1~226 - 4) である。ここには、プロセスをプロセッサ に固定にするかどうかの複報が入る。固定として おくとそのプロセスは、プロセッサ戦別子321に ロマカス

施3は、グループ機別子326(326 - 1 ~ 326 - 4)である。ここには、そのプロセスがどのグループに属しているかを示す情報が入る。グループの機別子とグループ管理テーブルは1対1に対応する。

84 は、CPU時間である。これは、プロセス が使用したCPU時間であり、優先度の計算に使 用する。ただし、このCPU時間は1秒毎に1/ 4にしているので正確なCPU時間ではない。

新5 に、プロセスの変行順を決定する低免度 328~380である。低免疫が各プロセスに複数ある のは、プロセッサがに優免度を持っているためで ある。この個免度の数値が小さいほど、プロセス の優先成は主くなり、他の優先次の数値が大きい プロセスに優免して実行されなければならないこ とを示す。

グループ管理テーブルには、プロセッや機関チ \$31 (321-1~331-2) がある。これは、その グループに属するプロセスを実行するプロセッサ 番号が記されている。

第9 関は、第8 医におけるスケジューラブログラム314のアルゴリズムを示している。スケジューラブログラムを実行しているプロセッやは、プロセスを選チーブルの優先度のうちむ分に対応するものを関係に関べて行く。そして、優先度が最大(係先度の数字は最小)のプロセスを選択する(314-1)。ただし、このとをプロセスを選択するプル32のプロセスの第27 でなっており、プロセス説別子321 が自分のプロセッサ番号でないプロセスは強く。

次に選択したプロセスを含分のプロセッサに割 り当てる (314 - 2)。

餅10回の315はタイマー14 - 2が発生をせる10

- 39 -

mo 包の割り込みの処窓ルーチンである。ここでは、 各プロセッサが実行中のプロセスのプロセス管理 テーブル32のCP U時間327をインクリメントし

第11回の316はタイマー14 - 2が発生させる1 = ことの前り込みの無限ルーチンである。ここでは、 会プロセス管理テーブル37の低先度328 ~ 830を所計算している。316 - 1 でもプロセスについて処理を行なっており、316 - 3 できプロセッサについて処理を行なっている。316 - 2 では、CPU 時間327が大きくなり過ぎないように1 秒毎に1 ノイにしている。

318-4~316-7で優先度を計算している。優先度は、基本的にはそのプロセスのCPU時間に受しいが、プロセス管理テーブルS2のプロセック度割子321に配きれていないプロセックの優先度には、20を加える。また、グループ管理テーブル33のプロセック機別子に書かれていないプロセックの優先度にはやは926を加える。

これにより、プロセス管理テーブルやグループ

管理テーブルのプロセッサ級別子に書かれている プロセッサに対してはそのプロセスの係外度が高 くなり、実行されやすくなる。

. - 40 -

第12回は、シングルプロセッサ用に開発したプログラムを効果のスケジューリング方式のマルチプロセッサで実行した場合を示している。第12回において、12-1~12-2はプロセス、41-1~41-2は割り込み禁止処理、42-1~42-2は装銀使用処理、43-1~43-2は割り込み許可処理である。

シングルプロセッサは、一幅一瞬を見ると実行 しているプロセスは一つである。よって、シング ルプロセッサで実行することを前掛としたプログ ラムは、共有製器を使用する場合にも、単に倒り 込みを額止するだけで使行しているものが多い。 これは、製御をロックして使用するより処理が高 速なためである。シングルプロセッサでは朝りな みさえ黄止しておけば、瞬間のには実行されて るプロセスは一つなので製紙を同時に複数のプロ セスが使用するようなことはない。

—435—

铃朋平 3-113563(12)

しかし、マルチプロセッサでは第12回に示すように、41-1,41-2で割り込みを禁止しても、 42-1と42-2のように同時に複数のプロセスが 登録を報信することがある。

#13回は、本発明のスケジューリング方法を用いているマルチプロセッサでシングルプロセッサ 用のプログラムを実行した場合を示す。何じ受数を使用するプロセス12-1と12-2を同一グループにして、プロセッサので実行するようにすれば、42-1と42-2の変数の使用が必ず直列に実行され、変慮使用が総合することがない。

尚、近年、個々にプログラムとデータを持つ従来のプロセスに代わって、いくつか低にプログラムやデータを共有するスレッド (thread) やライト・ウエイト・プロセス(light veight process)が注目されている。

プログラムやデータを共有するスレッドでは共 有データ量が多いため、グループ化したときのキャッシュ性態の向上が一段と大きくなる。

またプログラム、データを共有するスレッドを

グループ化して、1つのプロセッサで実行すると、 プログラム。データの切り替えがいらないことが 多くなり、OSの処理時間を知くすることができ

従って、本発明は、かかるスレッドやライト・ ウエイト・プロをスに対して適用しても効果大で ***

(器項の効果)

本意切のプロセススケジューリング方法は、複数のプロセッサからなるマルチプロセッサシステムにおいて、複数のプロセッサのいずれでも実行可能なプロセスについて、あえてそのプロセスをグループ化してそのグループを実行させるプロセッサを特定のプロセッサに領定しようとするものである。

共有データを多く持つプロセスが異なるプロセッサで実行されると、キャッシュ一致化のための 強備が頻繁に発売しキャッシュ性能を低下させる。 本売明では、そのようなプロセスをグルーブ化し、 回一のプロセッサで表行できるので、キャッシュ

٠ 43 -

一般なのための通信が不受となり、性能を脚上させることができる。さらに、グループを特定のプロセッサに変わさせることによりプロセッサ間の 気質が片寄ることがないように、一定時間以上実行されずに放置されているプロセスについては、いずれのプロセッサでも実行できるようにしており、気荷分数が可能となっている。ただし、この機能はプロセス固定ブラグを立てることにより禁止することもできる。

また、本発明によって、単色制質を必要とするプロセス容を何一のプロセッタで動作をせることが可能となり、共有メモリに対する徘徊制御をおり口となり、共有メモリに対する徘徊制御をおけるしていない位案の単一プロセッタ用の既存のプロセスを修正することができる。この事は、他のプロセッタの不使用を意味するものではなく、前記プロセス群以外で、任意のプロセッタで表行可信ととかできる。

- 48 -

- 44 -

4、図面の簡単な説明

第1個は本発明の一実施例のプロセススケジュ ーリングの概要回、第2回はシステム構成団、第 3 関はスケジューラのアルゴリズム領、戦4 優は スケジューラアルゴリズムの別の実施例図、第5 回もスケジューラのアルゴリズムの町の実動例図。 48.8 凶は選択したプロセスを自分のプロセッサに 割り当てるアルゴリズム図、第7回はタイマー割 り込みの処理ルーチン団、獲8回はシステム構成 の別の尖胞側図、第9団はスケジューラのアルゴ リズム区、第10回はloms年のタイマー割り込みの 処理ルーチン図、第11個は1秒経のタイヤー割り 込みの英雄ルーテン関、第12回は健康のスケジュ ーリング方式でシングルプロセッサ用のプログラ ムセ実行した場合の説明画、終18届は本意明のス ケジューリング方位でシングルプロセッサ用のプ ログラムを実行した聯合の設慰因である。

4 …接続バス、5 …共有メモリ、11 …プロセッ サ番号、12 …プロセス、13 … プロセッサ、14 … タ イマー、15 … グループ、31 … 〇 8 プログラム、32

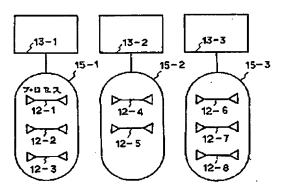
特期平 3-113563(13)

…プロセス管理テーブル、33…グループ管理チーブル、511…スケジューラ、321…プロセッサ級別子、822…優先成、323…夜行時も時間、324…カウンタ、325…プロセッサ国定フラグ、325…グループ機別子、327…CPU時間、328~330…優先度、331…プロセッサ説別子、332…女行時も時間、333…カウンタ、534…グループ固定フラグ、41… 製り込み禁止是選、42…安類使用処理、43…割り込み許可必無。

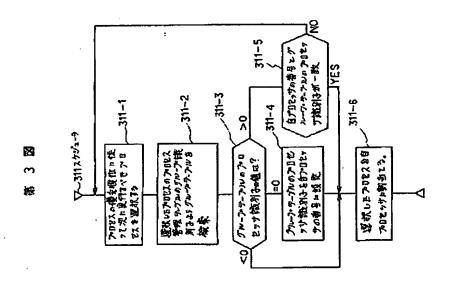
代理人 杂通士 秋 本 正 美

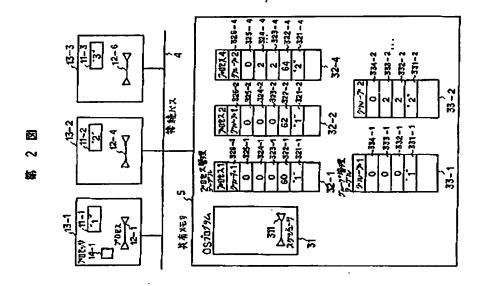
- 47 -

第 1 図

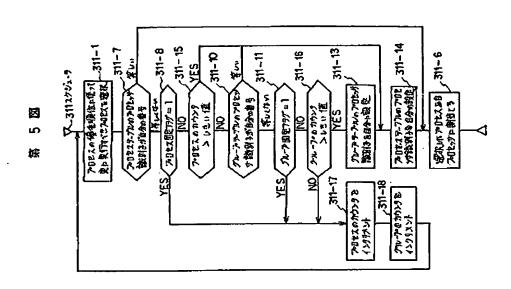


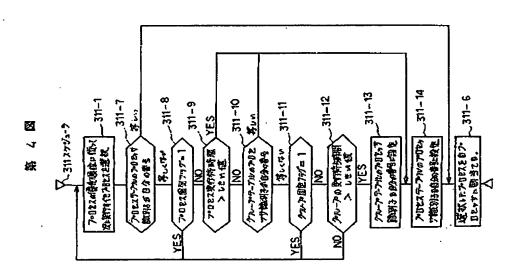
♦開平 3-113563(14)



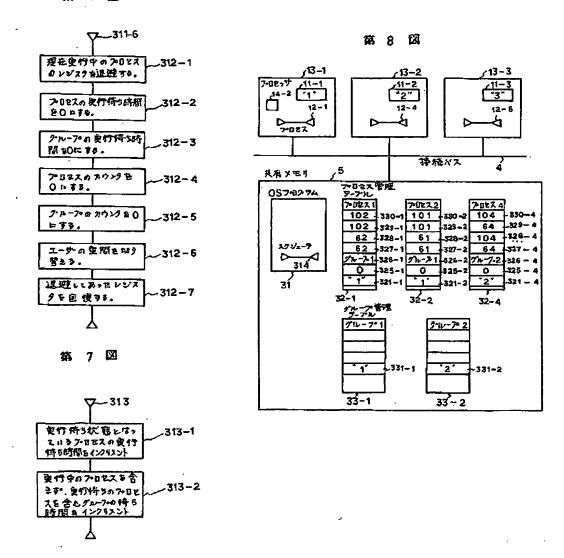


特朗平 3-113563(15)

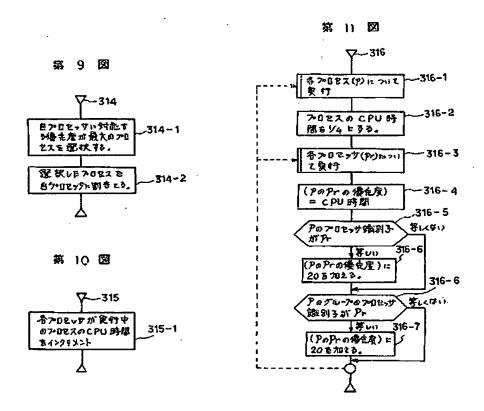




第 6 図

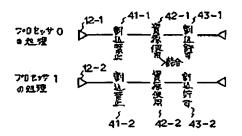


铃赛平 8-113563(17)

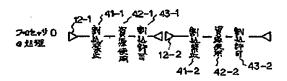


特別平 3-113563(18)

部 12 図



舊 13 図



第1頁の統令 ②発 明 者 中 村 智 明 茨城県日立市大みか町5丁目2番1号 株式会社日立製作 所大みか工場内